

POLICY RESEARCH WORKING PAPER

4484

# On Measuring the Benefits of Lower Transport Costs

*Hanan G. Jacoby*

*Bart Minten*

The World Bank  
Development Research Group  
Sustainable Rural and Urban Development Team  
January 2008



## Abstract

Despite large amounts invested in rural roads in developing countries, little is known about their benefits. This paper derives an expression for the willingness-to-pay for a reduction in transport costs from the canonical agricultural household model and uses it to estimate the benefits of a hypothetical road project. Estimation is based on novel cross-sectional data collected in a small region of Madagascar with enormous, yet plausibly

exogenous, variation in transport cost. A road that essentially eliminated transport costs in the study area would boost the incomes of the remotest households—those facing transport costs of about \$75/ton—by nearly half, mostly by raising non-farm earnings. This benefit estimate is contrasted to one based on a hedonic approach.

---

This paper—a product of the Sustainable Rural and Urban Development Team, Development Research Group—is part of a larger effort in the department to understand the poverty impacts of rural infrastructure. Policy Research Working Papers are also posted on the Web at <http://econ.worldbank.org>. The author may be contacted at [hjacob@worldbank.org](mailto:hjacoby@worldbank.org).

*The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.*

# On Measuring the Benefits of Lower Transport Costs

Hanan G. Jacoby\* and Bart Minten†

December 2007

---

\*Corresponding Author: Research Department, The World Bank, 1818 H Street NW, Washington DC, 20433 (e-mail: [hjacob@worldbank.org](mailto:hjacoby@worldbank.org); phone: 202-320-8666; fax: 202-522-1151).

†International Food Policy Research Institute, New Delhi (e-mail: [b.minten@cgiar.org](mailto:b.minten@cgiar.org)).

# 1 Introduction

Transport infrastructure investment in developing countries has received renewed attention in the past decade with greater recognition of the links between high transport costs and poverty (World Bank, 2007). Yet, despite large amounts spent on rural roads, remarkably little formal evidence exists on their benefits at the household level. What has been lacking is a general methodology for estimating these gains using micro-data.<sup>1</sup> Measuring the *benefit* of a road improvement involves more than estimating its various *impacts*. Any meaningful measure of welfare benefits must be grounded in a coherent economic model of the household; in the rural context, the agricultural household. Only by specifying how a particular road improvement, whether actual or hypothetical, would change the parameters of such a model, are welfare statements possible.

With its huge portfolio of transport-sector loans, the World Bank has historically taken the lead in developing appraisal techniques for rural road projects.<sup>2</sup> A recommended method for measuring benefits is to estimate the savings in transport costs that the project would entail, imputing the opportunity cost of travel time to market for households that do their own transport (see Lebo and Schelling 2001). However, in a setting where most households purchase transport services indirectly by accepting lower prices for their outputs (i.e., by selling to collectors at the farmgate or at intermediate points; see Fafchamps and Hill, 2005), this strategy may seriously understate the demand for transport. Moreover, a full accounting of the change in surplus due to the road project must include the gains from the additional economic activity it would generate.

This paper develops benefit measures that are attentive to both these issues. We use the canonical agricultural model to derive an expression for the willingness-to-pay for a reduction in transport costs; the area under the appropriately defined demand curve for transport. We also show how this benefit measure is related to one based on a hedonic approach. Both the ‘direct’ and hedonic benefits can be estimated nonparametrically using information that is, for the most part, commonly available in household survey data.

The central empirical obstacle to implementing our measures, and to estimating road impacts more generally, is reverse causation: Roads are not randomly placed (nor could they be feasibly randomized by design), and people do not randomly settle next to roads once they have been constructed. The causal link between better road access and the benefits of such access may thus be hopelessly obscured. Longitudinal data spanning a period of road construction can alleviate the endogenous road placement problem insofar

---

<sup>1</sup>An exception is Jacoby (2000), who uses a hedonic approach, similar to one derived in this paper, but based on more specialized assumptions, to estimate road project benefits in Nepal.

<sup>2</sup>van der Tak and Ray (1971) provide the classic statement of the social surplus approach based on supply and demand elasticities, although its implementation on micro-data would not be straightforward and, to our knowledge, has not been attempted.

as the unobservables determining such placement are fixed over time.<sup>3</sup> However, the time-frame for such evaluations may be too short to capture the long-term adaptations to lower transport costs (see also Mu and van de Walle, 2007). Moreover, collecting such data is very time-consuming and expensive, especially given the sample sizes needed to detect the often subtle impacts of road rehabilitation.

This paper takes a different tack. We use a one-shot cross-sectional data set with a novel sampling scheme. Households were surveyed in a small, relatively homogeneous region of rural Madagascar, over which transport costs to the *same* market vary tremendously. This variation is not due to the particularities of road placement – there are no paved roads and little motorized transport to speak of – but rather to the impenetrable mountains that range up and down the region. At their peak, transport costs amount to a staggering 50% of the final market price of the primary commodity, paddy rice. Arguably, a comparison of household behavior along this steep transport cost gradient approximates the long-run adjustments to an exogenous road improvement.

Nevertheless, to assess the validity of causal inferences about the various impacts of transport costs drawn from such cross-sectional data, we carry out two tests prior to our main estimation. The first test, presented in section 2, looks for systematic differences in land productivity across space. Such differences might explain why roads are not built into certain areas, and thus why these areas remain remote. At the same time, location-specific productivity may well be correlated with household level outcomes, such as land values, farm input use and output. The second test is for the presence of systematic differences in farm productivity across households at a given location; in particular, between migrant and native households. Poorer, less able migrants may be attracted to more remote areas by virtue of cheaper land. This selective migration could make the impacts of inaccessibility appear greater than they really are.

Before turning to these tests, section 2 describes our data and sampling methodology in detail. This section also presents a set of stylized facts that undergird our agricultural household model. Section 3 lays out the model and derives formulae for the willingness-to-pay for a reduction in transport costs. In Section 4, we implement these formulae using a nonparametric estimation procedure. Section 5 recaps the results.

---

<sup>3</sup>To the extent that road improvements are planned for areas with high expected growth, this assumption may be invalid. Only a data set that follows the same households over time would be useful for dealing with selective migration to areas with better road access. Khandker et al. (2006) use such data to estimate various impacts of road construction in Bangladesh, but they do not provide a comprehensive benefit measure.

## 2 Data and Background

### 2.1 Setting and sampling

Madagascar is a poor, rugged, and transport-deficient island-nation of Sub-Saharan Africa. After long years of neglect, the government, supported by international donors, has recently begun to spend substantially to rehabilitate and expand its road network.

The setting for our study is approximately 50 square kilometers of the Madagascar highlands, southeast of one of the largest towns in the country, Antsirabe. Households were surveyed over a broad backward L-shaped swath of countryside from Betafo, a town just west of Antsirabe, southwards to the environs of Bemaha above the Mania river, which practically seals off the region from the south,<sup>4</sup> and then eastwards to the isolated village of Andrembesoa and the valleys to its south and east (see figure 1).

**Figure 1: Map of Study Area**

For sampling purposes, the study area was divided into three zones according to remoteness. The least remote zone, the Betafo-Bemaha axis, consists of four communes contributing 900 households, or half the sample. Within each commune, five *fokontany* (extended villages) were randomly selected excluding the principal village of the commune (*chef de commune*). Then, within each of these *fokontany*, five hamlets were randomly chosen and a household list drawn up. Nine households were randomly selected from this list for interview. The second and third zones consist of the remote commune of Andrembesoa with its five *fokontany* (excluding the *chef de commune*). Four of these *fokontany*, make up our second zone, in which half of all households in each hamlet were randomly sampled (640 in all). Finally, *fokontany* Lohandany, nestled in the most westward and isolated valley on the map, makes up zone three, in which every one of the 258 households were surveyed.

Although the most distant households in our sample are less than 60 km from Antsirabe as the crow flies, the walk would take a local person about 3 days. Access to the main north-south highway in the east is blocked by mountain ranges (see figure 1). Moreover, due to the extremely steep terrain, much of the region, outside of the main valleys, is inaccessible to wheeled transport. Thus, agricultural output must generally be headloaded to the larger villages (e.g., Bemaha or Andrembesoa) for trans-shipment by ox-cart to the vicinity of Antsirabe.

Transport cost data are obtained as follows: Key informants in each of the 101 hamlets were asked about all markets utilized by the inhabitants for selling output, which in many cases were only intermediate points of sale on the way to Antsirabe. For each market,

---

<sup>4</sup>The Mania has no bridges (except on the national highway far to the east) and it is difficult and/or expensive to organize a crossing by pirogue.

information was gathered on the price per kilogram of porter services (inclusive of in-kind payments) and of ox-cart transport, depending on which was available. If both modes were available, we always took the ox-cart price; human portage is around 25% more expensive per hour of travel time than is transport by ox-cart. Transport costs are also somewhat higher during the wet season, but we use prices in the dry season when most transport occurs. For hamlets that do not ship directly to Antsirabe, we construct total costs per kilogram to the final market by summing the costs of each leg of the minimum cost route thereto.

Figure 2 shows the distribution of the 1761 sampled households according to transport costs to Antsirabe. The bimodality reflects the oversampling of households in the most remote areas; our sample is *not* designed to be representative of the spatial distribution of the region’s population. At any rate, for almost half of our sample, total transport costs lie in the range of 140-180 ariary per kilogram (2100 ariary = 1 USD) or around 75 USD per ton. By comparison, during harvest season 2006 (April-June), paddy rice sold for 300-350 ariary/kg in Antsirabe. Nonetheless, as we will see, many of these remote households still sell rice for export to Antsirabe.

### **Figure 2: Distribution of Sample Households by Transport Costs**

The data collection was split into two rounds to deal with seasonality issues. The first round survey was done in November-December 2005 with a 5-6 month recall to the previous July, and the second round took place in July-August 2006, covering the period from the preceding round to the end of June. Thus, overall, the survey has a 12 month recall from July 2005-June 2006, with main season agricultural production (principally rice) recorded in the second round and off-season production (often vegetables) recorded in the first round. Sales of rice and other staples occurring in July, for example, are from stocks held over from the 2005 harvest, whereas, starting in April, these sales come from the 2006 harvest. In the first round, 1798 households were interviewed, 37 of which, for various reasons, could not be followed up in the second round.

## **2.2 How heterogeneous is land productivity across space?**

Our first task is to assess the exogeneity of transport costs with respect to the household behaviors and outcomes of interest. This can be done indirectly, by comparing land productivity across our study area. We have information on rice yield for about 3300 lowland plots cultivated by nearly 1700 households. There are, however, two confounding factors to consider: First, input use is likely to decline with transport costs. That there are indeed important differences in production practices by location is shown in Figure 3. Fertilizers and other modern techniques, such as improved transplanting and weeding methods, are much less likely to be used on more remote rice plots. A second issue is

that, during the growing season for which we collected yield data, rainfall was relatively poor and especially so in the more remote part of our study area.

### Figure 3: Use of Agricultural Techniques by Transport Costs

Separating these influences requires estimating a production function for rice. Let log yield (assuming constant returns to scale) on plot  $p$  of household  $h$  depend on a vector of inputs  $x_{hp}$  according to

$$\log(y/a)_{hp} = f(x_{hp}) + \xi_{hp} + \nu_h + \epsilon_{hp} \quad (1)$$

where  $\xi_{hp}$  is an observed plot-specific weather shock and the error terms  $\nu_h$  and  $\epsilon_{hp}$  represent, respectively, household and plot-specific unobservables. We assume that  $Ex_{hp}\epsilon_{hp} = 0$ , but that  $x_{hp}$  and  $\nu_h$  are not necessarily mutually orthogonal. Since there are, on average, two rice plots per household, we can purge  $\nu_h$  using household fixed effects to obtain consistent estimates of the production function parameters. After estimating equation 1, we then calculate the residual

$$\hat{\eta}_{hp} = \exp(\overline{\log(y/a)} + \log(y/a)_{hp} - \hat{f}(x_{hp}) - \hat{\xi}_{hp}), \quad (2)$$

which is yield net of variation in weather shocks and input use. Finally, we estimate the function  $\hat{\eta}(\tau)$  nonparametrically, where  $\tau$  is transport costs as shown in figure 2.

Figure 4 summarizes the results of these procedures (see Appendix for the production function estimates). The solid curve shows the nonparametric regression of raw rice yield against  $\tau$ . The decline in yield is fairly substantial, from around 27 to 20 kg/are, though not monotonic. Adjusting yields for plot-level weather shocks alone (rainfall worse/much worse than normal; flooding worse than normal) changes the picture quite dramatically, resulting in somewhat of a u-shape pattern. While rainfall was generally poor during the 2005-06 rice growing season, self-reported drought conditions were much more prevalent on remote plots.<sup>5</sup>

### Figure 4: Rice Yield and Transport Costs

The dashed curve in Figure 4 removes the influence of input use on yields to get at underlying land productivity. This has the effect of flattening out the yield gradient to a considerable extent. If anything, rice productivity rises slightly with transport costs, although this tendency is not statistically significant in a linear regression of  $\hat{\eta}_{hp}$  on  $\tau$ . Thus, the inaccessibility of the most remote corner of our study area does not appear

---

<sup>5</sup>Splitting our sample of rice plots according to whether  $\tau$  is above or below the sample median, we find that the incidence of worse than normal drought conditions is 65% on remote plots and only 27% on non-remote plots.



to be the *result* of declining land productivity as one approaches it, because underlying land productivity does not, in fact, decline. This is not to imply that yield, rather than total output, is the relevant productivity indicator as far as road placement decisions are concerned. Total rice output of a region depends on yields as well as on the amount of available lowlands, a dimension along which the narrow valleys around Andrembesoa are not well endowed. Our point here is that productivity *per unit land area* is much more directly tied to the *household* outcomes of interest, and hence more germane to the endogeneity issue, than is the total land under cultivation in the region.

## 2.3 Migration and migrant selectivity

We now consider the endogeneity or selection issues arising from human settlement patterns. A striking feature of our study area is the extent of in-migration. We use a loose definition of migration for the moment, requiring only that the household head was not born in the *fokontany* of current residence. Thus, a child brought into the present location by his parents several decades ago and now a head of household is still considered a migrant. This designation yields 477 migrant-headed households compared to 1284 non-migrant. These migrants are largely Merina, the main highland ethnic group, but also include Betsileo from the neighboring Fianarantsoa province. About half of these migrant household heads had arrived within the past 15 years.

Figure 5 traces out a dramatic increase in the proportion of migrants with transport costs; 82% of migrant households in our sample reside in the most remote commune, Andrembesoa. Evidently, these migrants have been attracted by cheap and/or available agricultural land rather than by the ‘bright lights’ of the main town, Antsirabe. The salient question, for our purposes, is whether migrant and non-migrant households differ substantially. In terms of education, primary school completion rates of household heads are about equally low; 17% among non-migrants versus 16% among migrants. Similarly, differences in land ownership are trivial: Migrant households possess marginally less land than non-migrants (83.5 versus 86.3 are), whereas they own slightly more of the valuable lowland (34.3 versus 33.1 are).

### Figure 5: Proportion of Migrant Households and Transport Costs

Next, we ask whether migrants are selected on the basis of unobservables that are relevant to agricultural production. To do so, we calculate a measure of household farming ability based on the fixed effect,  $\nu_h$ , from the rice production function of equation 1. Since our previous, more general, analysis shows that  $\hat{\eta}'(\tau)$  is approximately zero, we can already rule out the case where migrant farmers are inferior to non-migrant farmers and the productivity of a given type is constant over space; had this story been true, then

$\hat{\eta}'(\tau)$  should have been negative. The question we can address here, therefore, is the extent to which migrants are *positively* selected on the basis of farming ability.

In Table 1, we regress the estimated fixed effect on household migrant status and a full set of hamlet dummies, which purge  $\tau$ , leaving us with a comparison of average farm productivity between migrant and native households at each location. Specification (1) indicates that migrant households are indeed positively, albeit weakly, selected, having nearly 7% *higher* farm productivity (in terms of rice output per unit area) than non-migrant households; this difference is not quite significant at the 5% level. A similar result obtains after controlling for other household characteristics in specification (2). It is also worth asking whether the selection is greater for more recent migrants. So, we construct a new indicator for household migration based on whether the head was older than age 18 at the time of the migration. One would expect that someone who migrated as an adult would be less similar to a native farmer than someone who migrated as a child and later started his own farm, and hence that the productivity differential between native and migrant farmers thus defined would be accentuated. However, specifications (3) and (4) of Table 1 show that, if anything, the opposite seems to be the case. There is essentially no productivity difference between households whose heads migrated as adults (18% of the sample) and those households that are either native or whose heads migrated as children. This result casts doubt on any selection, positive or negative, of households into remote areas on the basis of farming skills.

**Table 1: Migration and Rice Productivity**

Since there are few migrants living around Antsirabe, we are less concerned about selective migration into peri-urban areas on the basis of commercial orientation, entrepreneurial ability, or unobserved labor market skills. Insofar as these attributes are correlated with farming skills, however, lack of selection in one suggests lack of selection in the other. In section 4, we further explore this issue by disaggregating the off-farm earnings patterns of migrants and non-migrants.

## 2.4 Stylized facts

To motivate our model, we now present some stylized facts about the study area. As is typical of the Madagascar highlands, most crop production by our sampled households is for subsistence with relatively little sold for cash. Crop sales that do occur tend to conform to the classic von Thünen pattern of decreasing bulkiness with distance to market (as also noted by Jacoby, 2000, and Fafchamps and Shilpi, 2003, both in the context of rural Nepal). For example, among the least remote households ( $\tau < 40$  ariary/kg), 87% of crop sales by weight are either fruits, vegetables, or tubers, whereas among the most

remote households ( $\tau \geq 160$  ariary/kg), 95% of crop sales by weight are high value per kilogram dry grains (mainly rice, maize, and beans).<sup>6</sup>

Figure 6 conveys the key features of the market for rice, the main staple. Two-thirds of annual rice consumption is out of own-stocks held over from the previous harvest. The balance of consumption, purchased during the lean season, comes largely from sources within the village (*fokontany*), especially in the more remote areas. This is not to say that many remote households *sell* rice to their fellow villagers during the lean season. Indeed, as the figure shows, the bulk of the marketed rice surplus in these areas is exported outside the *fokontany* during the few months following the main harvest. Very few households—presumably, only those with sufficient access to credit—would be in a position to hold stocks throughout the lean season to meet local consumption demand.<sup>7</sup> Households residing in close proximity to Antsirabe, however, have ready access to rice brought in from other regions of Madagascar (or from other countries) during the lean season. Much of their purchased rice is obtained from outside the *fokontany* (presumably from shops near town) and that which is purchased from within the *fokontany* mostly comes from local shops rather than from other households, as is the case in more remote areas.<sup>8</sup>

#### Figure 6: Rice Consumption, Purchases, and Sales by Transport Costs

Landholdings are generally quite small, with average operated area of less than a hectare, divided about 40:60 between irrigated lowlands for rice cultivation and rainfed uplands. As indicated in Figure 7, land tenancy is extremely limited. Although most land is acquired through inheritance, a substantial fraction of rice land is purchased, whereas uplands are more likely to have been originally cleared by the owner, especially in the more remote areas where there has been much in-migration. Figure 7 also shows that the average area of rice under cultivation increases with transport costs. Evidently, remote households are more reliant on rice as a cash crop, and, as already seen, purchase a smaller fraction of their rice consumption.

#### Figure 7: Cultivated Area by Transport Cost and Mode of Acquisition

Agricultural labor is the main form of off-farm employment, done by 68% of households in the sample; only 23% engaged in any non-agricultural employment. Overwhelmingly,

---

<sup>6</sup>We exclude fresh cassava sales from these calculations, which, because of their bulk, take place almost exclusively on the local market. Thus, our focus here is on ‘exports’.

<sup>7</sup>Credit constraints appear to be the only plausible explanation for the peculiar behavior of rice-exporting households: They sell rice just after harvest when the price is low to meet immediate consumption needs and purchase rice in the lean season when the price is high (after exhausting their stocks) from wages earned during the intervening period.

<sup>8</sup>Ordering the five communes in our survey by increasing remoteness, the proportion of households whose within-*fokontany* rice purchases were usually obtained from shops/traders as opposed to other households is 0.90, 0.85, 0.05, 0.27, and 0.16.

the market for agricultural labor is local. Looking across all agricultural jobs reported in the employment section of the questionnaire, 97% took place in the *fokontany* of residence, as compared to 70% of all non-agricultural jobs. As will be shown in section 4, non-agricultural earnings decline rapidly with remoteness. This is true of both wage earnings and net revenue from family enterprises, which, overall, comprises 63% of off-farm income.

### 3 Willingness-to-Pay for Transport Cost Reduction

The precise question we address in this section is how to measure welfare gains from a road project that effectuates a non-marginal reduction in transport costs between a market town and its surrounding hinterland. We will assume that the area served by the potential road is ‘small’ relative to the national goods market, wherein the market town is perfectly integrated. This qualifier puts us squarely within a partial equilibrium framework, in which goods prices in the market town are fixed. Thus, the approach developed here may not apply to improvements in a trunk road or in a road network that could affect prices along its entire extent. Our calculations would be useful, however, in evaluating a nationwide rural road rehabilitation program that improves many, widely dispersed, feeder routes.

#### 3.1 Basic framework: Short-run equilibrium

Consider a population of agricultural households facing a variable transport cost of  $\tau$  ariary per kilogram of freight shipped to or from the final goods market of Antsirabe. For convenience, we index a household’s location in space by  $\tau$ , but there are many households at each location. The exogenous distribution of land ownership  $A$  at location  $\tau$  is given by the cdf  $F(A)$ . We suppress the locational dependence of this distribution and assume, without loss of generality, that mean land ownership  $\bar{A}$  is constant, which is tantamount to assuming uniform population density across space. Later, in examining the long-run equilibrium, we relax the assumption of fixed population density. Households can also vary in their labor endowment  $T$ , which has distribution function  $G(T)$ .

Households cultivate a vector of crops  $Y = \{y_1, \dots, y_k\}$  according to the concave production functions  $y_i = f^i(a_i, l_i, q_i)$ , where  $a_i$ ,  $l_i$ , and  $q_i$  are, respectively, the land area, labor, and market purchased input (e.g., fertilizer) devoted to crop  $i$ . At Antsirabe, crop prices are given by the vector  $P^x = \{p_1^x, \dots, p_k^x\}$ , so that the effective price facing sellers of good  $i$  at location  $\tau$  is  $\tilde{p}_i^x = p_i^x - \tau$ .<sup>9</sup> Likewise, since the total weight of the input  $q = \sum q_i$  must be imported from the market town, its effective price is  $\tilde{v} = v + \tau$ .

---

<sup>9</sup>The model thus allows for substitution into less bulky crops as transport costs rise. This can be seen by examining the first-order condition,  $\frac{p_i^x - \tau}{p_j^x - \tau} = \frac{\partial f^j / \partial a_j}{\partial f^i / \partial a_i}$ , which shows that as  $\tau$  increases the marginal product of land allocated to the relatively high value per unit weight crop (crop  $i$  if  $p_i^x > p_j^x$ ) must be driven downward.

Both farm labor and land rental markets exist. We assume that family labor  $l_i^f$  and hired labor  $h_i$  are perfect substitutes in each production function; i.e.,  $l_i = l_i^f + h_i$ . Ignoring leisure, each household can allocate  $T$  units of labor between own farm production and work on other farms, as long as these other farms are also located at  $\tau$ . Denote off-farm agricultural labor by  $l^o = T - \sum l_i^f$  when  $\sum l_i^f < T$ . Since households cannot hire labor in and out simultaneously,  $h \cdot l^o = 0$ , where  $h = \sum h_i$ .<sup>10</sup>

Given the above assumptions, household income is

$$\iota = \tilde{P}^x \cdot Y - \tilde{v} q - w(\tau)(l - T) - r(\tau)(a - A) \quad (3)$$

where  $l = \sum l_i$  and  $a = \sum a_i$ ,  $w(\tau)$  is the local wage and  $r(\tau)$  land rent. Equilibrium in the land market requires that optimal cultivated area equal average land ownership,  $a(r, w, \tau) = \bar{A}$ ; labor market equilibrium requires that  $l(r, w, \tau) = \bar{T}$ . These two conditions determine how rents and wages vary across space in the short-run (in which  $\bar{A}$  is fixed).

Households are assumed to have convex preferences, represented by the utility function  $u = U(X, C, Z)$ , over three types of consumption goods:<sup>11</sup> (i) self-produced agricultural goods  $X = \{x_1, \dots, x_k\}$ ; (ii) imported bulk goods  $C = \{c_1, \dots, c_m\}$ ; and (iii) non-imported goods or non-bulky imports, which, for convenience, we represent by the scalar  $Z$  with price normalized to one. By definition, goods of type (i) and (iii) do not involve transport from the final market, whereas the household does incur a transport cost for type (ii) goods; thus,  $\tilde{P}^c = \{p_1^c + \tau, \dots, p_j^c + \tau\}$ . In the present context, type (ii) goods could be items such as rice during the lean season and other foodstuffs (cooking oil, fish) that are not locally produced. Type (iii) goods would include locally purchased staples (e.g., cassava, charcoal) as well as essentially weightless imported items such as cigarettes, tea, salt, and clothes.

With the separability built into the model, the household can be viewed as first choosing  $\{a, l, q, h\}$  to maximize household income (note that  $l_i^f$  and  $h_i$  are not separately determined when  $h_i > 0$ , but this is not relevant for anything that we do). In the second stage, households maximize  $U(X, C, Z)$  subject to  $\tilde{P}^x \cdot X + \tilde{P}^c \cdot C + Z = \iota^*(\tilde{P}^x, \tilde{v}, w(\tau), r(\tau), A, T)$ , yielding indirect utility  $V(\tilde{P}^x, \tilde{P}^c, \iota^*(\tilde{P}^x, \tilde{v}, w(\tau), r(\tau), A, T))$ .

Now let  $\mu(\tau, A, T)$  be the income compensation required make a household with endowment  $(A, T)$  and facing the prices, wage, and rent prevailing at location  $\tau$  indifferent to the situation prevailing at  $\tau = 0$ . In other words,  $\mu(\tau, A, T)$  is the equivalent variation

---

<sup>10</sup>We do not explicitly accounted for the time households spend taking their harvest to market. While output is often sold to collectors at the farmgate, much is also headloaded by household members to intermediate collection points for onward shipment to Antsirabe. In this latter case, we make the simplifying approximation that their opportunity cost of time is identical to the market rate for portage, which we observe (when relevant) at each location.

<sup>11</sup>Earlier transport cost models of Walters (1968) and Gersovitz (1989), albeit not directly concerned with measuring road benefits, consider only the production side of the farm household, whereas here we take into account the consumption side as well.

of a reduction of transport costs from  $\tau$  to zero and is implicitly defined by the identity

$$V(\tilde{P}^x, \tilde{P}^c, \iota^*(\tilde{P}^x, \tilde{v}, w(\tau), r(\tau), A, T) + \mu) \equiv V(P^x, P^c, \iota^*(P^x, v, w(0), r(0), A, T)). \quad (4)$$

By the envelope theorem  $\iota_\tau^* = -\sum_i y_i - q - w_\tau(\tau)(l - T) - r_\tau(\tau)(a - A)$  and by Roy's identity  $x_i = -V_{p_i^x}/V_\iota$  and  $c_j = -V_{p_j^c}/V_\iota$ , where subscripts denote partial derivatives. With these results in hand, we see that differentiation of equation 4 with respect to  $\tau$  and rearranging leads to

$$\mu_\tau(\tau, A, T) = \sum_j c_j + \sum_i (y_i - x_i) + q + w_\tau(\tau)(l - T) + r_\tau(\tau)(a - A). \quad (5)$$

The appearance of a simple aggregate of the weight of freight,  $s(\tau) \equiv \sum_j c_j + \sum_i (y_i - x_i) + q$ , on the right-hand side of equation 5 is reminiscent of the composite commodity theorem with  $s(\tau)$  acting as the 'demand' for transport tonnage and  $\tau$  as its price.<sup>12</sup> The demand for transport has a production component  $\sum_i y_i + q$ , which only depends on effective output and factor prices, and a consumption component  $\sigma(\tau) \equiv \sum_j c_j - \sum_i x_i$ , which depends on both goods prices and income.

Equation 5 represents a partial differential equation in  $\mu$ . Solving it for each  $(A, T)$  combination is complicated by the dependence of  $\sigma$  on  $\iota^* + \mu$ . However, if the income effect is zero, the benefit of a non-marginal reduction in transport costs from a baseline of  $\tau_0$  to zero can be had merely by integrating both sides to obtain

$$\mu(\tau_0, A, T) = \int_0^{\tau_0} [s(\tau) + w_\tau(\tau)(l - T) + r_\tau(\tau)(a - A)] d\tau. \quad (6)$$

As our interest is in the average benefit across households, we must also integrate over the endowment distributions,  $F$  and  $G$ , which can be done before the integration over  $\tau$  in equation 6. Since  $s$  does not depend on  $A$  and  $T$  – a consequence of both separability between consumption and production decisions and the assumption of zero income effects – we need only invoke the land and labor market equilibrium conditions to get

$$E[\mu(\tau_0)] = \int_0^{\tau_0} s(\tau) d\tau. \quad (7)$$

Thus, because the average household at each location is neither a net renter of land nor

---

<sup>12</sup>For households that do all of their own transport,  $s(\tau)$  could be replaced by, say, the number of journeys to the market, with  $\tau$  then being measured in terms of forgone earnings per journey. The advantage of our formulation, noted earlier, is its greater generality, being useful in settings where households sell at the farmgate or at intermediate collection points.

hirer of labor, the effects of changes in transport costs on wages and rents vanish on average.

Our benefit measure is then the area under the demand curve for transport, the units of which are in money (i.e., kg×ariary/kg). The assumption of zero income effects in consumption is tantamount to using the uncompensated (Marshallian) demand instead of the compensated (Hicksian) demand to calculate this consumer surplus.<sup>13</sup> In principle, it is possible to allow for income effects in  $\sigma$ , but estimating them reliably would be difficult because it requires exogenous variation in income that is independent of transport costs.<sup>14</sup>

Let us compare  $E[\mu(\tau_0)]$  to an alternative estimate of the benefits of lower transport costs that has been proposed implicitly, if not explicitly, in the literature. Consider, as in, e.g., Gibson and Rozelle (2003) or Khandker et al., (2006), a regression of household consumption expenditures on  $\tau$  (or on some other indicator of remoteness). How does the predicted change in average consumption from  $\tau_0$  to zero correspond to  $E[\mu(\tau_0)]$ ? Suppose, in particular, that we evaluate consumption,  $\widetilde{M}$ , at effective prices  $\widetilde{P}^x$  and  $\widetilde{P}^c$ , which is, roughly speaking, the conventional method of constructing expenditure aggregates. Now, differentiate the budget constraint,  $\widetilde{M} = \iota^*$ , with respect to  $\tau$  (continuing to ignore income effects in  $\sigma$ ), and integrate both sides over  $F$  and  $G$  to get  $E[\widetilde{M}_\tau] = \sigma - s$ . Using equation 7, it then follows that

$$E[\widetilde{M}(0)] - E[\widetilde{M}(\tau_0)] = E[\mu(\tau_0)] - \int_0^{\tau_0} \sigma(\tau) d\tau. \quad (8)$$

If, on average,  $\sigma < 0$ , which is a reasonable assumption in our data (see figure 6 for the case of rice), then the difference in average consumption expenditures *overestimates*  $E[\mu(\tau_0)]$ ; this is because  $E[\widetilde{M}(0)] - E[\widetilde{M}(\tau_0)]$  does not take into account substitution effects in consumption.

To capture the greater non-agricultural labor opportunities found near towns we must, however, augment benefit formula 7. As noted by Fafchamps and Shilpi (2005), these opportunities may arise from von Thünen-type transport cost economies or from agglomeration externalities (increasing returns to specialization with market size). Ideally, we would also like an explicit equilibrium model of non-agricultural labor earnings that would allow us to predict how these earnings change with a reduction in transport costs. However, not only would such a model require us to take a specific stand on which of the complex and poorly understood mechanisms underlies non-agricultural employment in developing countries, but it would also not be estimable given the structure of the current data set (i.e., effectively one town). We will have to settle, then, for a ‘reduced-form’

---

<sup>13</sup>It also implies that Marshallian consumer surplus, equivalent variation, and compensating variation all coincide. If the assumption is false, then  $E[\mu(\tau_0)]$  is bounded below by the equivalent variation and bounded above by the compensating variation (see, e.g., Hausman, 1981).

<sup>14</sup>Even with a good estimate of  $\partial\sigma/\partial\iota$ , solving differential equation 5 is far from trivial, as a practical matter. For one thing,  $\iota$  itself also varies with  $\tau$ .

approach in this case.

For simplicity, we assume that non-agricultural employment is rationed at a fixed wage, as in Harris and Todaro (1970) (see Fafchamps and Shilpi, 2005, for some evidence supporting such model of urban employment in Nepal). Without this assumption, we would have to explain why households with access to both agricultural and non-agricultural off-farm employment choose to do both, which is quite a common situation in the data. At any rate, letting non-agricultural earnings,  $e(\tau)$ , depend (negatively) on transport costs captures the higher probability of finding or being able to partake of such work near the main town. Our revised benefit formula thus becomes

$$E[\mu(\tau_0)] = \int_0^{\tau_0} s(\tau) d\tau + E[e(0)] - E[e(\tau_0)]. \quad (9)$$

### 3.2 A hedonic approach: Long-run equilibrium

An alternative measure of road benefits uses the hedonic approach based on land values (see also Jacoby, 2000). For the land value or rent gradient to reflect spatial welfare differentials requires that households be able to freely migrate across locations.<sup>15</sup> In this case, from the *ex ante* perspective of households deciding where to locate, the land endowment  $A$  is zero. At the same time, however, in-migration entails an *ex post* fall in  $\bar{A}$ , or, equivalently, a rise in population density. Since the marginal household must be indifferent to where it settles and buys or rents land, we have

$$V(\tilde{P}^x, \tilde{P}^c, \iota^*(\tilde{P}^x, \tilde{v}, w(\bar{A}, \bar{T}, \tau), r(\bar{A}, \bar{T}, \tau), 0, \bar{T})) \equiv V(P^x, P^c, \iota^*(P^x, v, w(0), r(0), 0, \bar{T})). \quad (10)$$

Here we assume, for simplicity, that migrants have the average labor endowment  $T = \bar{T}$ , and we ignore non-farm employment for the same reason.

Equation 10 implicitly defines  $\bar{A}(\tau)$ , average landholdings in long-run equilibrium. Let  $\frac{dr}{d\tau} = r_\tau + \frac{dr}{dA} \frac{d\bar{A}}{d\tau}$  be the overall rent gradient, the sum of the short-run gradient introduced earlier and the long-run effect on rents of changes in population density. Differentiating the equilibrium condition with respect to  $\tau$  and using Roy's identity once again gives  $\iota_\tau^* + \sigma = 0$ , where for the marginal household  $\iota_\tau^* = -\sum_i y_i - q - a \frac{dr}{d\tau}$ . Rearranging, we get

$$-\frac{dr}{d\tau} a = s(\tau) = E[\mu_\tau(\tau)]. \quad (11)$$

Replacing  $a$  by  $\bar{A}(\tau)$  (they must be the same in equilibrium) and integrating then delivers

---

<sup>15</sup>By contrast, much of the urban economics literature treats cities as closed to migration so that land values do not necessarily reflect the differential benefits from amenities. On this point see Arnott and Stiglitz (1981), Polinsky and Shavell (1976), as well as the survey of Bartik and Smith (1987).



the hedonic benefit formula

$$E[\mu(\tau_0)] = - \int_0^{\tau_0} \bar{A}(\tau) \frac{dr}{d\tau}(\tau) d\tau. \quad (12)$$

Equation 12 says that in the long-run free-migration equilibrium, the overall rent gradient must reflect the short-run benefits of variation in transport costs, as derived above.<sup>16</sup>

The reason why short-run benefits as measured by formula 7 are the same as the long-run benefits given in formula 12 is that we are considering *average* benefits for the extant population rather than total benefits summed over the whole population.<sup>17</sup> Since migration induces an increase in the number of households wherever transport costs have been reduced (nonmarginally), the societal benefit to a road improvement has an additional long-run component (Scotchmer, 1986). However, to take into account the gains from the redistribution of population across space, one would have to know how total population varies with  $\tau$ , which is beyond the scope of the present investigation.

While formulae 7 and 12 should thus produce the same benefit estimates in the context of our model, in practice they might not. One reason is that differences in land rent across space may also reflect the benefits of amenities that are related to transport costs, such as access to social services and other facilities. We return to this point in the next section.

## 4 Empirical Results

Both of our proposed benefit measures involve integrals of a function  $g(\tau, \theta)$ , where  $\theta$  is a vector of household characteristics (endowments, etc.), which may or may not be observed. The first step is to estimate the conditional expectation function  $\hat{\varphi}(\tau) = E_\theta[g(\tau, \theta)|\tau]$  nonparametrically. This means that  $\hat{\varphi}$  is calculated at a set of  $K$  grid-points  $\{\tau_{(1)}, \dots, \tau_{(K)}\}$  along the interval  $[0, \tau_0]$  and the desired integral is estimated by

$$\int_0^{\tau_0} E_\theta[g(\tau, \theta)|\tau] d\tau = \sum_{k=2}^K \hat{\varphi}(\tau_{(k)}) (\tau_{(k)} - \tau_{(k-1)}). \quad (13)$$

---

<sup>16</sup>All of our previous results go through in a setting in which new land is being brought into cultivation ( $A$  endogenous), provided that land is not a free resource. In particular, let  $\psi(A)$  be the cost of clearing and making cultivable  $A$  units of land. Staying within our static framework, optimal landholdings  $A^*$  solves  $\psi'(A) = r$  and land market equilibrium requires that  $a(r, w, \tau) = A^*(r)$ . It is easily shown that neither our direct or hedonic benefit formula are affected by this change. If agricultural land could truly be cleared at zero marginal cost, then the gains from the expansion of cultivated area would have to be factored into any calculation of benefits of rural road improvements. Of course, the price of cropland would effectively be zero in this case, rendering moot the hedonic method based on land values.

<sup>17</sup>Note that the only alteration in equation 5 in the long-run is that  $r_\tau$  and  $w_\tau$  are replaced by the respective overall gradients,  $\frac{dr}{d\tau}$  and  $\frac{dw}{d\tau}$ . Since, on average, these terms do not enter the benefit calculation for the extant population, there is no difference in average benefits between the short-run and the long-run.

We bootstrap standard errors of our benefit measures, treating *fokontany* as strata to be resampled independently, thus preserving the original sample’s geographic distribution.<sup>18</sup>

By taking zero, or the minimum transport cost, as the lower limit of integration, these benefit calculations assume that the hypothetical road project renders the most remote location as accessible to the final goods market as the least remote location. Of course, no road project can literally accomplish this feat. However, given the compactness of our study area, a good road passing through the most remote hamlet, allowing regular truck transport by a number of competitive carriers, would go most of the way in this direction.

#### 4.1 Direct benefit estimates

The results of our benefit calculations are summarized in Table 2. We first use equation 7, based only on the weight of freight. The production side contribution to the integrand,  $s(\tau)$ , consists of the sales of crop (except cassava) and livestock products (principally milk) plus bulk input purchases (chemical fertilizer and seed).<sup>19</sup> On the consumption side, we have to decide which household purchases are imported from the final market ( $C$  goods) and which are obtained from local production ( $Z$  goods). Thus, any bulk item purchased outside the *fokontany* is assumed to be imported from the final market, even if bought in an intermediate market, whereas anything purchased from within the *fokontany* is assumed to be from local production. Exceptions to the latter rule are cooking oil, fish and flour, which when purchased from *fokontany* sources are, according to our data, usually obtained from shops rather than from other households. The case of rice is more complicated, for reasons already discussed in section 2.4. We assume that locally purchased rice is purely from local production in the three outlying communes but purely imported in the two close-in communes (cf. footnote 8). While by no means perfect, these rules for allocating purchases should provide a reasonable approximation to the weight of  $C$  goods.

Figure 8 presents a nonparametric estimate of  $s(\tau)$  along with its various subcomponents. All the transport cost gradients are indeed negative, though in the case of purchased farm inputs the decline with  $\tau$  is obscured by the overall low level of use.

#### Figure 8: Weight of Freight and Transport Costs

Using this estimate of  $s(\tau)$ , we find that a road improvement making the remotest

---

<sup>18</sup>Benefits are likely to be estimated substantially more precisely than  $g$ . To see why, take the case where  $g$  is linear in  $\tau$ . Suppose that in half the bootstrap samples the estimated slope of  $g$  is  $-a$  and in the other half it is  $-a/2$ ,  $a > 0$ , but that the two estimates of  $g$  always intersect at the midpoint of  $[0, \tau_0]$ . Thus, while the standard error of the slope is positive, the standard error of the relevant area under  $g$  is zero. More generally, variations in the slope of  $g$  will tend to wash out when calculating the integral of the function.

<sup>19</sup>Cassava sales are excluded, as are purchases of manure, because these virtually all take place in the local market and thus do not involve transport.

hamlet in our sample as accessible as the least remote—equivalent to a reduction in transport costs of about 75 USD/ton—would be worth around 99,000 ariary of annual income. More than a third of this benefit (37% to be precise) is due to the lower effective price of imported consumption – the  $C$  goods.

To put benefits into relative terms, we calculate a household consumption expenditure aggregate,  $\widetilde{M}$ , valuing consumption of own production using median sales prices in each commune. A nonparametric estimate of  $\widetilde{M}(\tau)$ , partialling out variation in household size using partially linear regression, is shown in Figure 9. Taking  $E[\widetilde{M}(\tau_{(K)})] = 584,000$  ariary as an estimate of the annual income of the most remote household, yields a benefit-income ratio, as shown in the first column of Table 2, of 17%.

**Table 2: Alternative Benefit Estimates of a Transport Cost Reduction**

Next, we incorporate non-farm earnings using the benefit measure given by equation 9. Earnings are calculated as the sum of all salaries from nonagricultural employment plus net revenue from family enterprises during the year. The nonparametric estimate of  $e(\tau)$  in Figure 9 (again adjusting for household size) confirms a rapid drop-off in non-farm opportunities with increased remoteness. The combined benefit-income ratio of the road improvement, in the second column of Table 2, is 52%. Surprisingly, changes in effective prices for bulky commodities account for only a third of the benefits of greater accessibility.

**Figure 9: Non-farm Earnings and Transport Costs**

Compare this latter benefit figure to one derived solely from household consumption expenditures; in theory, at least,  $100 \times (E[\widetilde{M}(\tau_{(0)})] / E[\widetilde{M}(\tau_{(K)})] - 1)$  should exceed 52% (cf. equation 8). Indeed, this is precisely what we find. The expenditure-based benefit estimate is 64% of income (with a standard error of around 5%), which is 23% higher than the one based on the area under the transport demand curve.

## 4.2 Robustness: Selective migration

We next explore the implications of heterogeneity in transport cost gradients across migrant and non-migrant households. To fix ideas, let  $m$  be an indicator for migrant household and  $z$  be a household-level outcome of interest. We may write

$$E[z|\tau] = E[z|\tau, m = 1] \Pr(m = 1|\tau) + E[z|\tau, m = 0] (1 - \Pr(m = 1|\tau)), \quad (14)$$

which expresses the unconditional transport cost function as a weighted sum of the conditional functions for migrant and non-migrant households, where the weight,  $\Pr(m = 1|\tau)$ , is increasing in  $\tau$  as shown in Figure 5. Clearly, if  $E[z|\tau, m = 0] > E[z|\tau, m = 1] \forall \tau$ ,

then  $E[z|\tau]$  is decreasing in  $\tau$  even if  $E[z|\tau, m = 0]$  and  $E[z|\tau, m = 1]$  are themselves both constant.<sup>20</sup> For example, if migrants purchase less rice than non-migrants, for whatever reason, then we would erroneously conclude that overall consumption imports decline with  $\tau$ , or that they decline faster than they should. To avoid any such fallacy of composition, we estimate separate transport cost functions, and corresponding benefit measures, for migrant and non-migrant households and then sum the two using the estimated migration probability at the maximum transport cost,  $\tau_{(K)}$ , as the weight.

Following this approach, we recalculate the direct benefit measures in the second row of Table 2. The results show that not accounting for household migrant status in this fashion does significantly overstate the benefits of transport cost reduction in our sample, at least at the 10% level (using a conventional t-test with the bootstrapped standard error of the difference; see fourth row of Table 2). The main reason for this appears to be that non-farm earnings fall somewhat more slowly with  $\tau$  for migrants than for non-migrants. However, if we redo the calculation with migrants defined as those households whose heads migrated as adults, the overall benefit estimate is quite similar to its unadjusted counterpart, at 47% of annual income. Hence, just as in Table 1, and rather counterintuitively, the evidence for migration selectivity is considerably weakened when migrants are categorized more strictly.

### 4.3 Hedonic benefit estimates

The hedonic benefit formula relies on the land rent-transport cost gradient. But, given the limited land rental market in our setting, it would be difficult to estimate a rent gradient with any precision. Instead, we use reported land values (see Jacoby, 2000, for a discussion and justification of this approach). To convert land values into an annualized flow that allows comparison with our other benefit measures as well as with yearly income or expenditures, we need an estimate of the discount rate  $\delta$ . For this we exploit the simple present value formula for per are rent,  $r = \delta V$ , where  $V$  is the reported value per are of the plot. In our sample of 333 rented plots, the median value of  $r/V$  is 0.11; a practically identical estimate of  $\delta$  emerges from a regression analysis on the same sample.<sup>21</sup>

We estimate separate nonparametric regressions of the land value gradient for lowland and upland plots. To deal with heterogeneity in land quality, we control for observed plot

---

<sup>20</sup>In the special case where the transport cost gradient is a constant  $\alpha$  and where  $E[z|\tau, m = 1] - E[z|\tau, m = 0] = \beta$ , the problem boils down to one in which the true regression model is  $z = \alpha\tau + \beta m + u$  and  $m$  is an omitted variable. In the general case, equation 14 can be viewed as a semiparametric switching regression model. Given lack of migrant selectivity on the basis of unobservable agricultural productivity (cf. Table 1), the difficulty of estimating endogenous switching models, and the exploratory nature of this analysis, we assume that the switching is exogenous.

<sup>21</sup>Specifically, we regress  $r$  (including the value of in-kind rent payments) on  $V$  (missing values imputed using information on type of plot, plot area, and  $\tau$ ) without a constant term (null of zero constant cannot be rejected with  $p$ -value=0.17). In this regression, we instrument  $V$  for measurement error using plot area. The coefficient on  $V$  is 0.113 with a standard error of 0.014.

characteristics using a partially linear regression. In the case of lowland plots, we have plot topography, soil type, type of irrigation, irrigation water availability, and frequency of flooding (see Appendix for details). For the rainfed upland plots, we only collected information on plot topography. The estimation sample consists of 2767 lowland plots and 2976 upland plots.

Figure 10 shows the nonparametric land value gradients. For lowlands, there appears to be a precipitous rise in land values within a relatively short distance of Antsirabe, followed by a steady decline as transport costs increase. Upland plots, which are of much lower average value than rice plots in general, do not display this marked initial rise. Our survey enumerators indicated that rice plots in the commune of Belazao, the one nearest to Antsirabe, were considered inferior to other lowlands in the region. Indeed, after dropping the 11% of lowland plots contributed by households from Belazao, the anomalous result largely disappears and the land value-transport cost relationship becomes practically monotonic. In our hedonic benefit calculations, therefore, we eliminate households from Belazao, there being still plenty of close-in households to spare in our sample.

### Figure 10: Value of Lowland and Upland Plots and Transport Costs

To calculate the integral in equation 12, we multiply the plot-type specific value gradient by the discount rate and the respective estimate of  $\bar{A}$  for each type of land, add these two numbers together at each grid-point, and then form the weighted sum over grid-points as shown in equation 13. The result of this calculation is reported in the last column of Table 2, and it is comparatively large. The hedonic benefit estimate for the most remote households is 607,000 ariary per year, or 104% of annual income.

Given the proportionality implied by the present value formula, however, the absolute level of benefits derived from any hedonic method based on land values will clearly be sensitive to assumptions about the discount rate. Jacoby (2000), for example, estimates a discount rate of only 6% in rural Nepal. Setting  $\delta = 0.06$  in the present case would practically equalize the hedonic and direct estimates of road benefits.

Aside from assumptions about the discount rate, the discrepancy between our hedonic and direct estimates could be due to locational amenities correlated with transport costs. As mentioned earlier, the hedonic benefit impounds the value of such amenities. One urban amenity that *cannot* reconcile the two benefit measures, though, is greater access to labor market opportunities; non-farm employment earnings are already incorporated into our direct benefit estimates.

What about differences in access to social services? Although primary schools are present in every *fokontany* in our study area, secondary schooling is only available near Antsirabe. Consequently, the secondary school enrollment rate for children 13-18 years-old is 31% among close-in households ( $\tau < 40$ ) compared to only 8% for remote households

( $\tau \geq 160$ ). Nevertheless, given such low overall enrollment, it seems difficult to explain the large gap between direct and hedonic benefits on this basis alone. As for medical care, there is little difference across locations in the proportion of households visiting a health clinic during the year.

Another advantage of being near town might be access to financial services, especially formal credit sources. But, again, there is very little formal credit reported in our data; only 25 households had any such debt in the two years preceding the survey and even these households are not heavily concentrated around Antsirabe. We can also ask about agricultural extension services. Although only 9% of households overall reported receiving an extension visit in the last five years, these are indeed much rarer in the more remote areas. However, in an analysis of rice productivity, similar to that done for migration in Table 1, we find no significant impact of extension visits on yields conditional on transport costs. Thus, it appears unlikely that a substantial component of our hedonic benefit estimate reflects the demand for social, financial, or extension services, although we certainly cannot rule out the possibility of some other, unobserved, urban amenities.

## 5 Conclusions

Evaluating the impact of transport infrastructure poses considerable empirical challenges. ‘Natural’ experiments involving large, permanent, changes in transport costs (and only in transport costs), for which the long-run equilibrium is observed both before and after the intervention, are indeed rare, if not nonexistent. In this paper, we have attempted to circumvent this problem by collecting cross-sectional data in a region over which, due to the vagaries of topography, the ratio of transport costs to the main staple price varies from 0.05 to 0.5. Analysis of spatial patterns in the productivity of this staple shows that, at least on this important dimension, the sample area is relatively homogeneous. While this may alleviate concerns about reverse causation, the possibility that the transport cost gradients estimated in this paper are contaminated by other unobservable correlates of accessibility can never be ruled out entirely. Putting aside this caveat, our chief interest here has not been to catalogue the multifarious impacts of changes in transport costs, as much as to measure the potential welfare benefits of rural roads within a coherent and generalizable economic framework.

Based on our preferred direct benefit estimate, a hypothetical rural road project that reduces the transport costs of the most remote households by around 75 USD per ton would raise their incomes by about 50 percent. But the gain due to the reduction in the cost of goods transport, both exports and imports, is small compared to that from improved access to non-farm earnings opportunities in town. This suggests that there may be potentially important complementarities between rural road construction and urban

economic development.

We also present results from a hedonic approach. Given the apparently limited relevance of locational amenities, this benefit estimate should be similar to that obtained from the direct approach. It turns out to be substantially larger, but the comparison between the two measures is heavily dependent on the chosen discount rate. Needless to say, replication of the methodology developed in this paper across different settings would enhance our understanding of both the size and composition of the benefits of rural roads.

We close with the observation that the benefits from road construction are generally not equally distributed; it is obviously the most remote households that have the most to gain from improved access. Even though land rents may adjust so as to dissipate the net benefits for incoming migrants, and drive the gains to zero at the margin, incumbent landowners are unambiguously better off from a road improvement. Thus, to the extent that remote households are poorer to begin with, a policy of building rural roads can have desirable distributional properties (a point elaborated on at length in Jacoby, 2000). A related issue, however, is that, by rewarding incumbents, road building may encourage too much migration into remote areas, whether in anticipation of a road being built or, especially where land for new settlement is still available, after a road project has been announced. The implications of this form of rent-seeking for public investments and spatial development are potentially far-reaching, but go well beyond the purview of the present paper.

## References

- [1] Arnott, R. and J. Stiglitz (1981): "Aggregate land rents and aggregate transport costs," *Economic Journal*, 91(June):331-47.
- [2] Bartik, T. and V. Smith, Urban amenities and public policy. In: E.S. Mills, Editor, *Handbook of Regional and Urban Economics* vol. 2, North-Holland, Amsterdam (1987), pp. 1207–1254.
- [3] Fafchamps, M. and F. Shilpi (2003): "The Spatial Division of Labour in Nepal," *Journal of Development Studies*, 39(6):23-66.
- [4] Fafchamps, M. and F. Shilpi (2005): "Cities and specialisation: evidence from South Asia," *Economic Journal*, 115(503): 477-504.
- [5] Fafchamps, M. and R. Hill (2005): "Selling at the farmgate or traveling to market," *American Journal of Agricultural Economics*, 87(3):717-34.
- [6] Gersovitz, M. (1989): "Transportation, state marketing, and the taxation of the agricultural hinterland," *Journal of Political Economy*, 97(5):1113-1137.
- [7] Gibson, J. and S. Rozelle (2003): "Poverty and access to roads in Papua New Guinea," *Economic Development and Cultural Change*, 52(1): 159-85.
- [8] Harris, J. and M. Todaro (1970): "Migration, unemployment and development: a two-sector analysis," *American Economic Review*, 60: 126-42.
- [9] Hausman, J. (1981): "Exact consumer's surplus and deadweight loss," *American Economic Review*, 71(4):662-76.
- [10] Jacoby, H. G. (2000): "Access to markets and the benefits of rural roads," *Economic Journal*, 110(465): 713-737.
- [11] Khandker, S., Z. Bakht and G. Koolwal (2006): "The poverty impact of rural roads: evidence from Bangladesh," unpublished manuscript, The World Bank, Washington, DC.
- [12] Lebo, J. and D. Schelling (2001): *Design and Appraisal of Rural Transport Infrastructure: Ensuring Basic Access for Rural Communities*. World Bank Technical Paper No. 496. Washington, DC: The World Bank.
- [13] Mu, R. and D. van de Walle (2007): "Rural Roads and Poor Area Development in Vietnam," Policy Research Working Paper No. 4340, Washington, DC: The World Bank.



- [14] Polinsky, A. M. and S. Shavell (1976): "Amenities and property values in a model of an urban area," *Journal of Public Economics*, 5:119-29.
- [15] Scotchmer, S. (1986): "The short-run and long-run benefits of environmental improvement," *Journal of Public Economics*, 30:61-81.
- [16] van der Tak, H. G. and A. Ray (1971): *The Economic Benefits of Road Transport Projects*. World Bank Staff Occasional Papers No. 13. Washington, DC: The World Bank.
- [17] Walters, A. (1968): *The Economics of Road User Charges*. World Bank Staff Occasional Papers No. 5. Washington, DC: The World Bank.
- [18] World Bank (2007): *A Decade of Action in Transport: An Evaluation of World Bank Assistance to the Transport Sector, 1995-2005*. Washington, DC: The World Bank.

**Table 1: Migration and Rice Productivity**

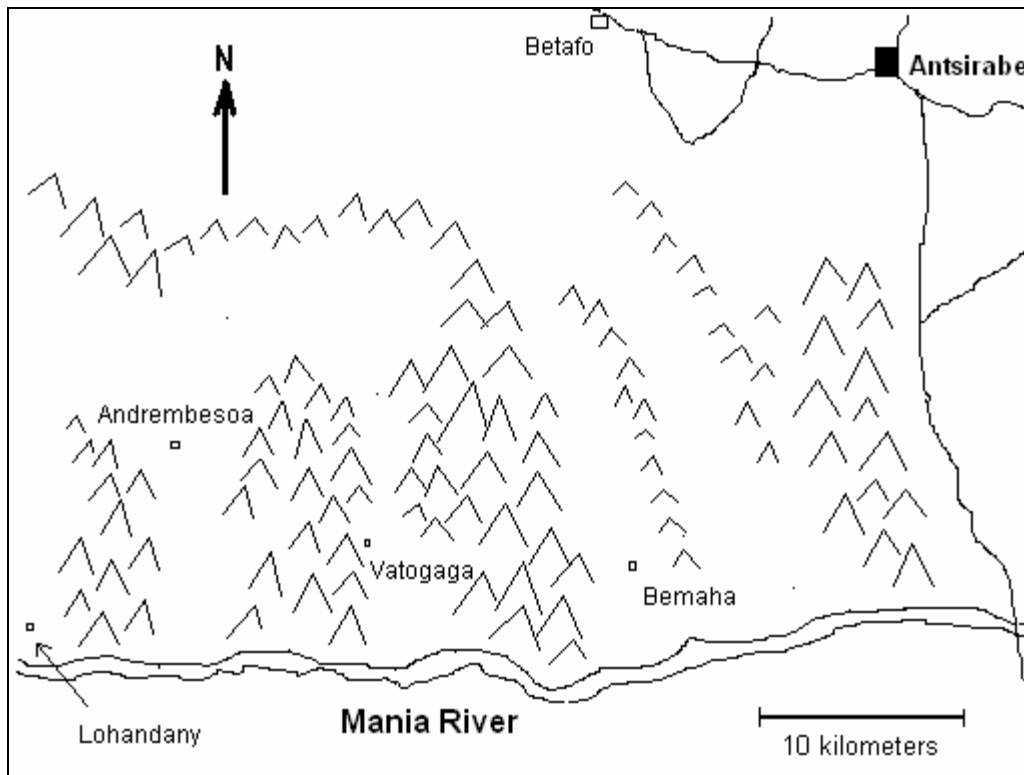
	<b>Means</b>	<b>(1)</b>	<b>(2)</b>	<b>(3)</b>	<b>(4)</b>
Head migrated	0.27	0.069 (0.038)	0.065 (0.036)	---	---
Head migrated as an adult (older than age 18)	0.18	---	---	0.019 (0.036)	0.013 (0.034)
Head completed primary school	0.16	---	-0.019 (0.034)	---	-0.016 (0.034)
Head is female	0.08	---	0.097 (0.057)	---	0.098 (0.056)
Head's age	40.8 (14.3)	---	0.005 (0.005)	---	0.005 (0.005)
Head's age squared		---	-0.000 (0.000)	---	-0.000 (0.000)
Number of household members older than 12	3.2 (1.6)	---	0.004 (0.008)	---	0.003 (0.008)

*Notes:* Robust standard errors (adjusted for clustering by hamlet) in parentheses. Dependent variable is the estimated household fixed effect from the rice production function (see Appendix). All specifications include hamlet fixed effects and use a sample of 1681 households.

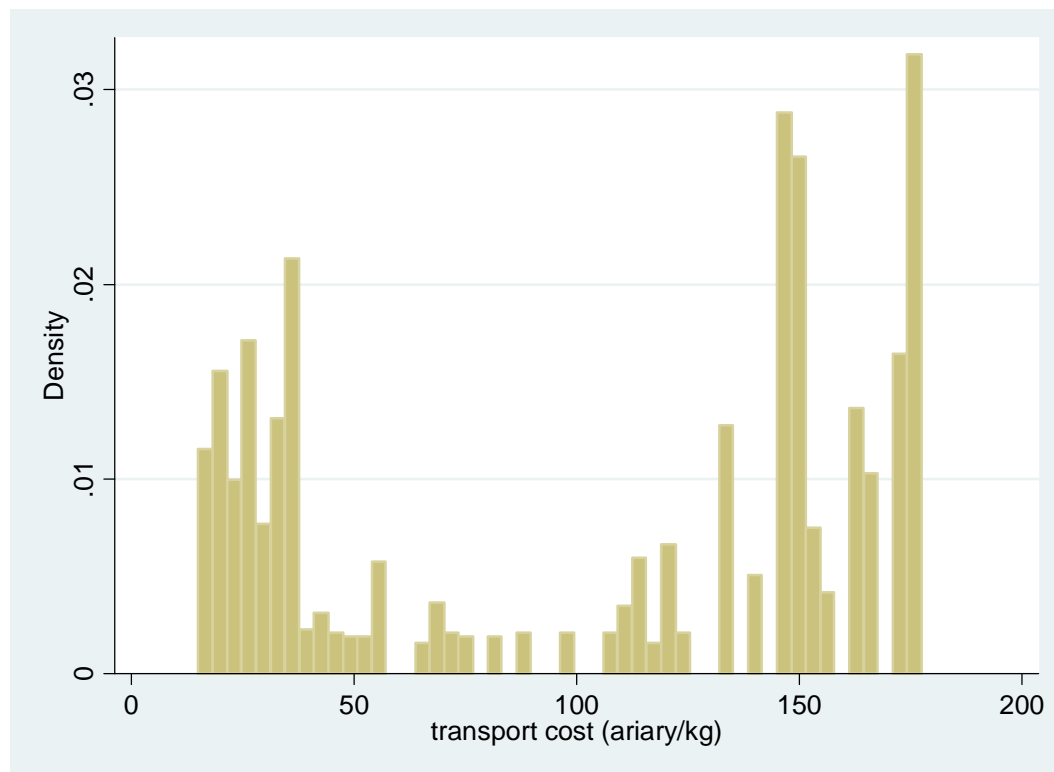
**Table 2: Alternative Benefit Estimates of a Transport Cost Reduction**

	<i>Benefit /income</i>		
	Freight only	Freight & non-farm earnings	Hedonic
(1) Unadjusted	0.170 (0.006)	0.518 (0.048)	1.039 (0.029)
(2) Adjusted for migration	0.143 (0.010)	0.422 (0.062)	---
(3) Adjusted for migration as adult	0.147 (0.008)	0.475 (0.064)	---
(1) – (2)	0.026 (0.009)	0.096 (0.051)	---
(1) – (3)	0.023 (0.007)	0.043 (0.047)	---

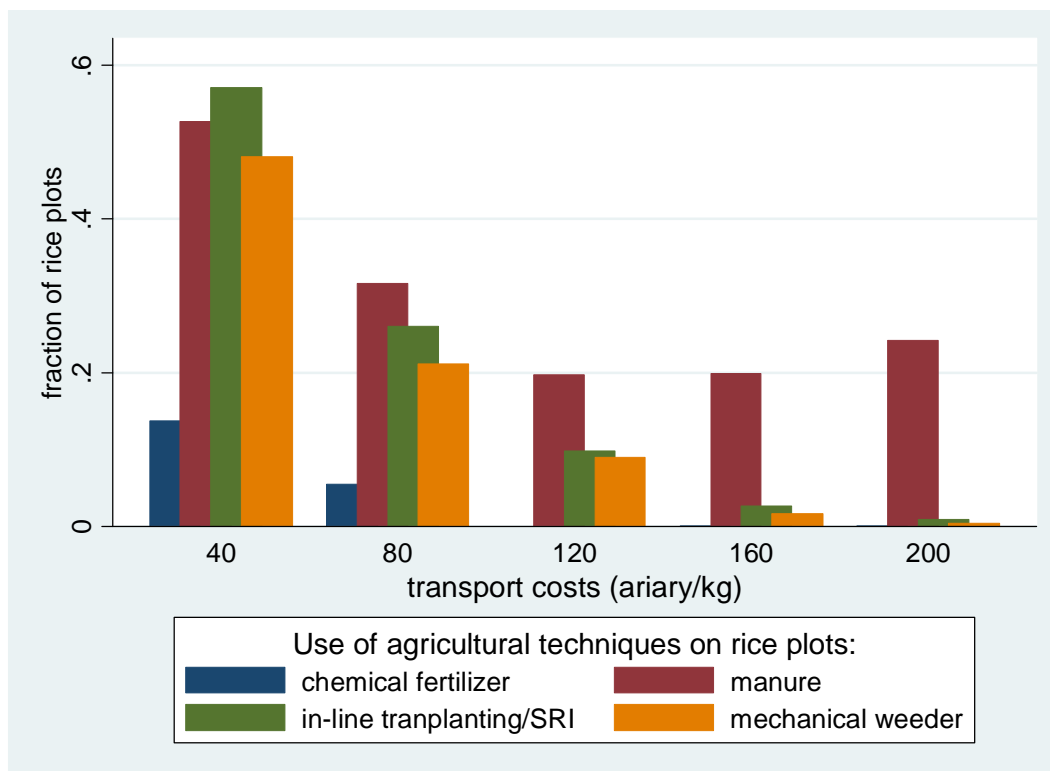
*Notes:* Bootstrapped standard errors (in parentheses) treat each of the 25 *fokontany* as a separate strata. Figures in the first, second, and third columns are based on formulae 7, 9, and 12, respectively. Benefit is defined as the compensating variation required to make households in the least remote location ( $\tau = 15$  ariary/kg) indifferent to residing in the most remote location ( $\tau = 178$  ariary/kg). Estimation of annual income for the most remote households is described in the text. For the hedonic measure, the least remote location has  $\tau = 22.5$  ariary/kg because all plots in the close-in commune of Belazao were dropped for reasons discussed in the text. Migration adjustment estimates benefits separately for migrant and non-migrant groups and takes the weighted sum of the results, with weights equal to the probability of being a migrant evaluated at the most remote location (see text).



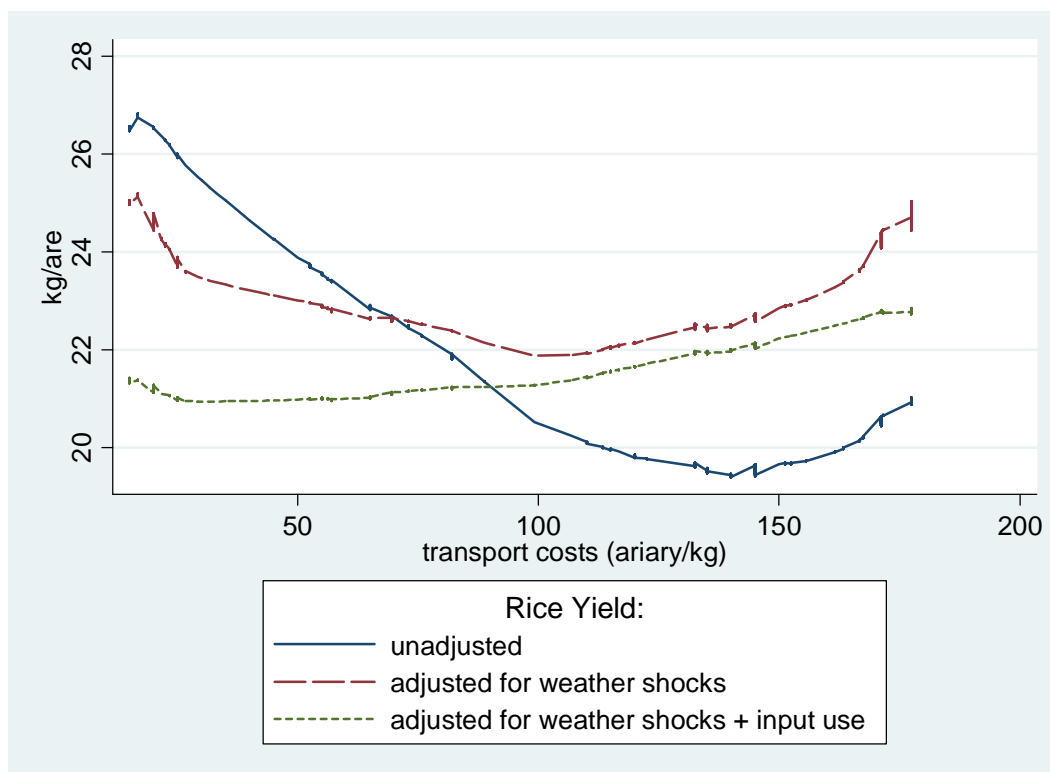
**Figure 1: Map of Study Area**



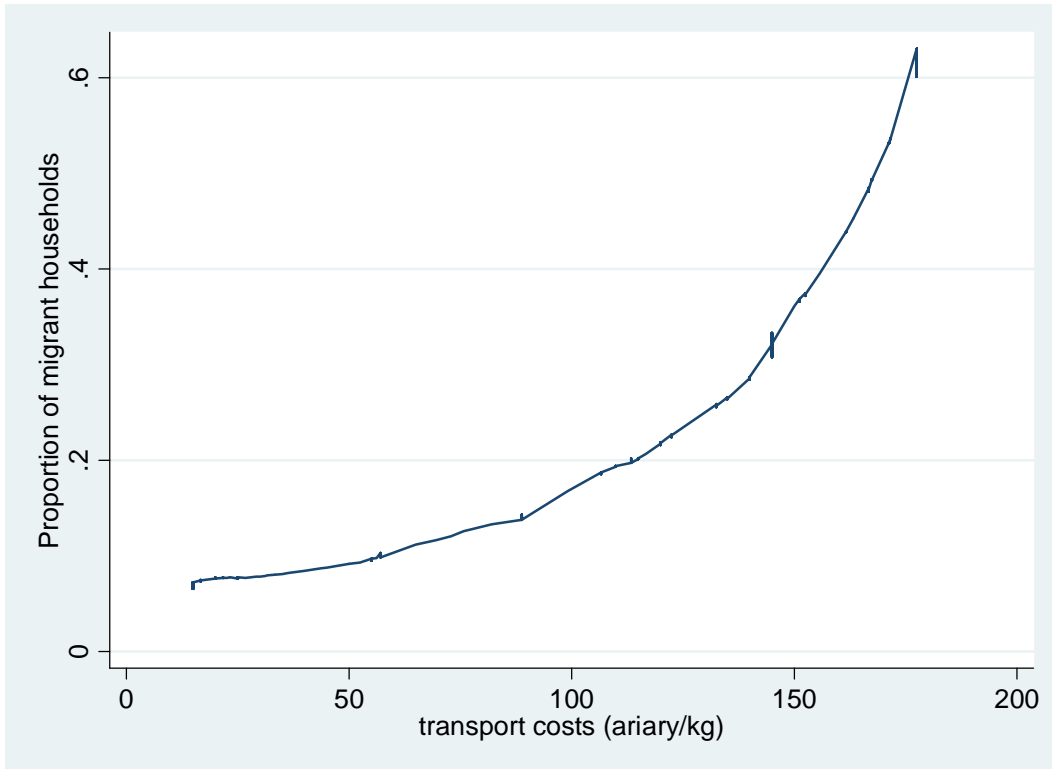
**Figure 2: Distribution of Sample Households by Transport Costs**



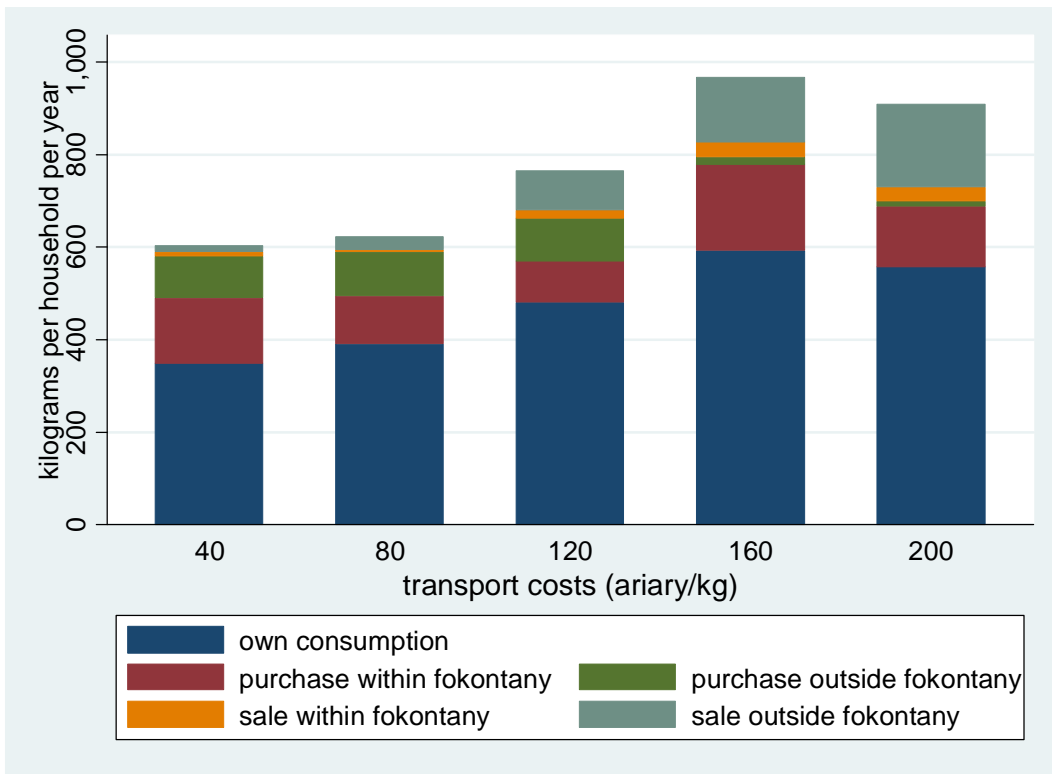
**Figure 3: Use of Agricultural Techniques by Transport Costs**



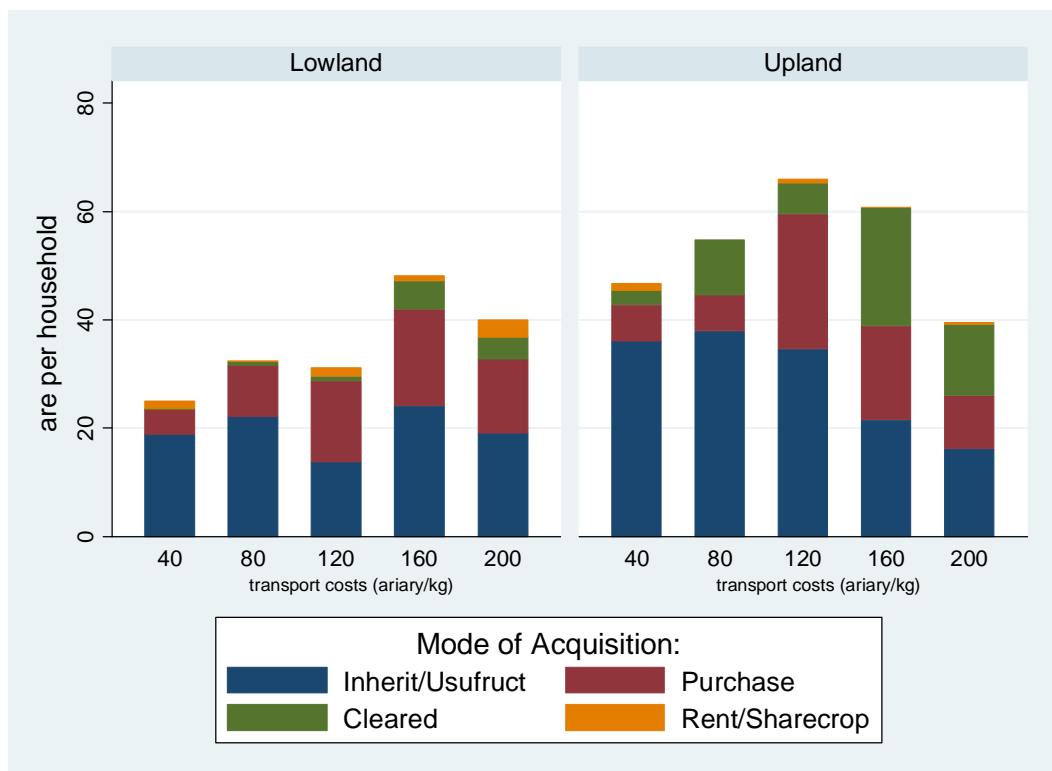
**Figure 4: Rice Yield and Transport Costs**



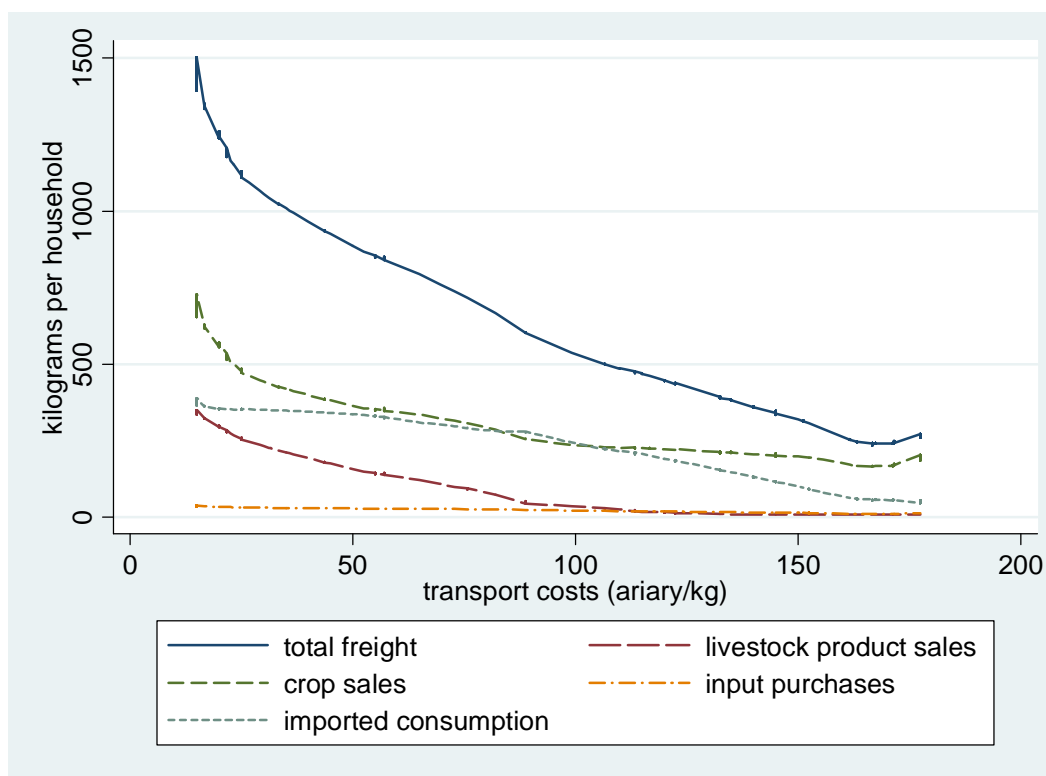
**Figure 5: Proportion of Migrant Households and Transport Costs**



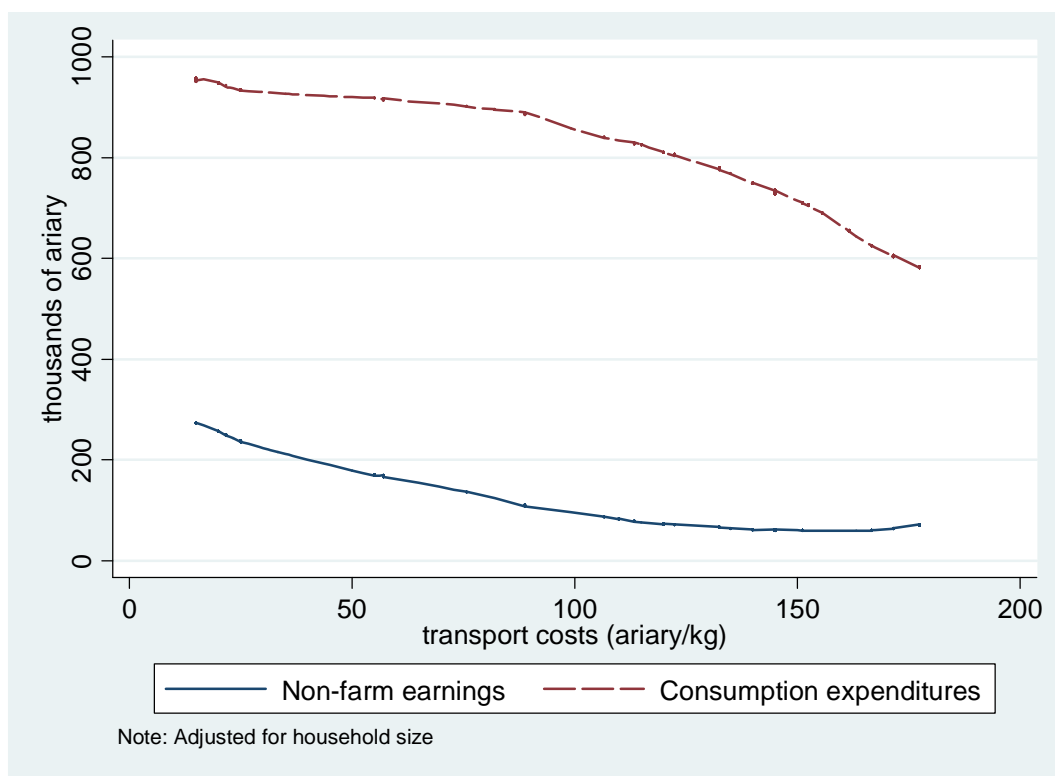
**Figure 6: Rice Consumption, Purchases, and Sales by Transport Costs**



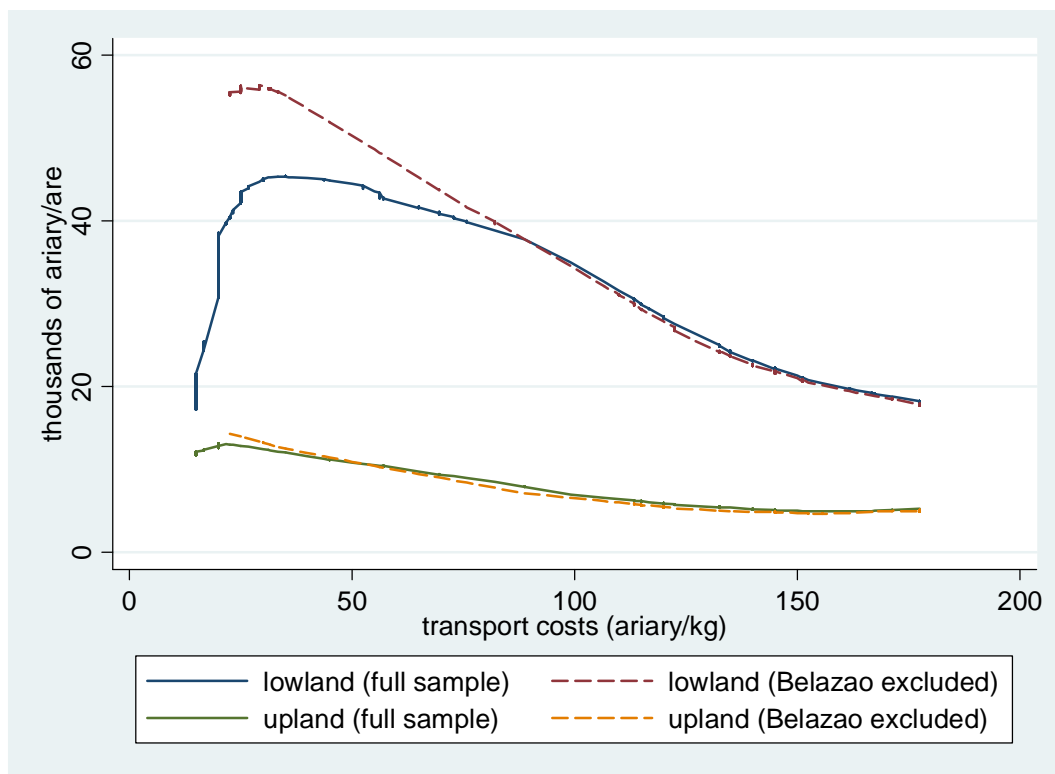
**Figure 7: Cultivated Area by Transport Costs and Mode of Acquisition**



**Figure 8: Weight of Freight and Transport Costs**



**Figure 9: Non-farm Earnings, Consumption, and Transport Costs**



**Figure 10: Value of Lowland and Upland Plots and Transport Costs**



### Appendix: Estimates of Rice Production Function

	(1)	(2)
drought much worse than expected	<b>-0.831</b> (0.057)	<b>-0.712</b> (0.056)
drought worse than expected	<b>-0.208</b> (0.037)	<b>-0.141</b> (0.037)
log(kg of seed/are + 1)	---	<b>0.398</b> (0.066)
any chemical fertilizer	---	<b>0.097</b> (0.066)
log(kg of manure/are + 1)	---	<b>0.008</b> (0.006)
log(hours of animal traction/are + 1)	---	<b>0.177</b> (0.065)
log(days of labor/are + 1)	---	<b>0.406</b> (0.051)
modern transplanting technique	---	<b>0.095</b> (0.059)
multiple weedings	---	<b>0.206</b> (0.057)
irrigation – traditional	---	<b>0.013</b> (0.034)
irrigation – source	---	<b>0.108</b> (0.040)
topography – base of hill	---	<b>-0.041</b> (0.025)
topography – terraced	---	<b>-0.003</b> (0.030)
black soil <sup>a</sup>	---	<b>0.077</b> (0.021)
water availability index <sup>a</sup>	---	<b>-0.056</b> (0.020)
flooding frequency index <sup>a</sup>	---	<b>-0.004</b> (0.013)
H <sub>0</sub> : random effects (p-value)	0.103	0.0000
Effective sample size	2647	2638

Notes: Standard errors in parentheses. All specifications include household fixed effects. Effective sample drops plots of households contributing only a single rice plot. Dependent variable is the log of rice yield (kg/are). **Details of selected independent variables:** Drought perception is plot-specific self-report (omitted category: drought normal or better than expected). Chemical fertilizer quantity not included because normal practice is to apply very small amounts to seedbed prior to transplanting. Modern transplanting is either in-line or SRI (mostly the former) as opposed to transplanting in bunches. Irrigation is either by traditional method, from a water source, or rainfed (omitted category). For topography, omitted category is bottom land. For soil, omitted category is red or brown. Water availability index takes on values 1-5, with 1 being always available and 5 being never available. Flooding frequency index takes on values 1-4, with 1 being frequent and 4 being never.

<sup>a</sup>These variables are treated as fixed plot characteristics, not subject to intervention by the farmer; hence they are not netted out of adjusted yield  $\hat{\eta}_{hp}$  (see equation (12) in text).